BIOTECHNOLOGY
SYSTEMS
BRANCH

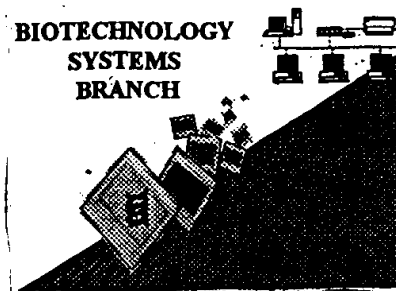# RAW SEQUENCE LISTING ERROR REPORT

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:

Application Serial Number: **10/016,634**
Source: **OIPE**
Date Processed by STIC: **12/31/2001**

THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.
PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:
1) **INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,**
2) **TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY**
FOR CRF SUBMISSION QUESTIONS, PLEASE CONTACT MARK SPENCER, 703-308-4212.

FOR SEQUENCE RULES INTERPRETATION, PLEASE CONTACT ROBERT WAX, 703-308-4216.
PATENTIN 2.1 e-mail help: patin21help@uspto.gov or phone 703-306-4119 (R. Wax)
PATENTIN 3.0 e-mail help: patin3help@uspto.gov or phone 703-306-4119 (R. Wax)

TO REDUCE ERRORED SEQUENCE LISTINGS, PLEASE USE THE **CHECKER VERSION 3.1 PROGRAM**, ACCESSIBLE THROUGH THE U.S. PATENT AND TRADEMARK OFFICE WEBSITE. SEE BELOW FOR ADDRESS:
    http://www.uspto.gov/web/offices/pac/checker

Applicants submitting genetic sequence information electronically on diskette or CD-Rom should be aware that there is a possibility that the disk/CD-Rom may have been affected by treatment given to all incoming mail.
Please consider using alternate methods of submission for the disk/CD-Rom or replacement disk/CD-Rom, including:
1. EFS-Bio (<**http://www.uspto.gov/ebc/efs/downloads/documents.htm**> , **EFS Submission User Manual** - ePAVE)
2. U.S. Patent and Trademark Office, Box Sequence, P.O. Box 2327, Arlington, VA 22202
3. **Hand Carry** directly to:
    U.S. Patent and Trademark Office, Technology Center 1600, Reception Area, 7th Floor, Examiner Name, Sequence Information, Crystal Mall One, 1911 South Clark Street, Arlington, VA 22202
            Or
    U.S. Patent and Trademark Office, Box Sequence, Customer Window, Lobby, Room 1B03, Crystal Plaza Two, 2011 South Clark Place, Arlington, VA 22202
4. **Federal Express, United Parcel Service, or other delivery service to:** U.S. Patent and Trademark Office, Box Sequence, Room 1B03-Mailroom, Crystal Plaza Two, 2011 South Clark Place, Arlington, VA 22202

ERROR DETECTED	SUGGESTED CORRECTION	SERIAL NUMBER: 10/016,634

ATTN: NEW RULES CASES: PLEASE DISREGARD ENGLISH "ALPHA" HEADERS, WHICH WERE INSERTED BY PTO SOFTWAF

1_____Wrapped Nucleics
Wrapped Aminos

The number/text at the end of each line "wrapped" down to the next line. This may occur if your file was retrieved in a word processor after creating it. Please adjust your right margin to .3; this will prevent "wrapping."

2_____Invalid Line Length

The rules require that a line not exceed 72 characters in length. This includes white spaces.

3_____Misaligned Amino Numbering

The numbering under each 5th amino acid is misaligned. Do not use tab codes between numbers; use space characters, instead.

4_____Non-ASCII

The submitted file was not saved in ASCII(DOS) text, as required by the Sequence Rules. Please ensure your subsequent submission is saved in ASCII text.

5_____Variable Length

Sequence(s)_____contain n's or Xaa's representing more than one residue. Per Sequence Rules, each n or Xaa can only represent a single residue. Please present the maximum number of each residue having variable length and indicate in the <220>-<223> section that some may be missing.

6_____PatentIn 2.0 "bug"

A "bug" in PatentIn version 2.0 has caused the <220>-<223> section to be missing from amino acid sequences(s) _____. Normally, PatentIn would automatically generate this section from the previously coded nucleic acid sequence. Please manually copy the relevant <220>-<223> section to the subsequent amino acid sequence. This applies to the mandatory <220>-<223> sections for Artificial or Unknown sequences.

7_____Skipped Sequences (OLD RULES)

Sequence(s) _____ missing. If intentional, please insert the following lines for each skipped sequence:
(2) INFORMATION FOR SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)
(i)   SEQUENCE CHARACTERISTICS: (Do not insert any subheadings under this heading)
(xi) SEQUENCE DESCRIPTION:SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)
This sequence is intentionally skipped

Please also adjust the "(ii) NUMBER OF SEQUENCES:" response to include the skipped sequences.

8_____Skipped Sequences (NEW RULES)

Sequence(s) _____ missing. If intentional, please insert the following lines for each skipped sequence.
<210> sequence id number
<400> sequence id number
000

9_____Use of n's or Xaa's (NEW RULES)

Use of n's and/or Xaa's have been detected in the Sequence Listing.
Per 1.823 of Sequence Rules, use of <220>-<223> is MANDATORY if n's or Xaa's are present.
In <220> to <223> section, please explain location of n or Xaa, and which residue n or Xaa represents.

10_____Invalid <213> Response

Per 1.823 of Sequence Rules, the only valid <213> responses are: Unknown, Artificial Sequence, or scientific name (Genus/species). <220>-<223> section is required when <213> response is Unknown or is Artificial Sequence

11_____Use of <220>

Sequence(s) _____ missing the <220> "Feature" and associated numeric identifiers and responses. Use of <220> to <223> is MANDATORY if <213> "Organism" response is "Artificial Sequence" or "Unknown." Please explain source of genetic material in <220> to <223> section.
(See "Federal Register," 06/01/1998, Vol. 63, No. 104, pp. 29631-32)   (Sec. 1.823 of Sequence Rules)

12_____PatentIn 2.0 "bug"

Please do not use "Copy to Disk" function of PatentIn version 2.0. This causes a corrupted file, resulting in missing mandatory numeric identifiers and responses (as indicated on raw sequence listing). Instead, please use "File Manager" or any other manual means to copy file to floppy disk.

13_____Misuse of n

n can only be used to represent a single nucleotide in a nucleic acid sequence. N is not used to represent any value not specifically a nucleotide.

AMC/MH – Biotechnology Systems Branch – 08/21/2001

OIPE

RAW SEQUENCE LISTING
PATENT APPLICATION: US/10/016,634

DATE: 12/31/2001
TIME: 14:11:01

Input Set : **A:\Dex-255.app**
Output Set: **N:\CRF3\12312001\J016634.raw**

Does Not Comply
Corrected Diskette Needed

```
   3 <110> APPLICANT: Sun, Yongming
   4         Recipon, Herve
   5         Ghosh, Malavika
   6         Liu, Chenghua
   8 <120> TITLE OF INVENTION: Compositions and Methods Relating to Colon Specific
   9         Genes and Proteins
  11 <130> FILE REFERENCE: DEX-0255
  13 <140> CURRENT APPLICATION NUMBER: US/10/016,634
  14 <141> CURRENT FILING DATE: 2001-10-31
  16 <150> PRIOR APPLICATION NUMBER: 60/244,758
  17 <151> PRIOR FILING DATE: 2000-10-31
  19 <160> NUMBER OF SEQ ID NOS: 176
  21 <170> SOFTWARE: PatentIn Ver. 2.1
```

ERRORED SEQUENCES

```
1354 <210> SEQ ID NO: 56
1355 <211> LENGTH: 271
1356 <212> TYPE: DNA
1357 <213> ORGANISM: Homo sapiens
1359 <400> SEQUENCE: 56
1360 catcttttta atattcagta tgaccgaata aagcactggt gctgccttag taacaatggt 60
E--> 1361 tgtnctcaag gtaaacttct catgtgcttg tttcagttgt gagctcaatt agcctctttc 120
1362 tcatgaaatg aatgcctttt tacttgaaag aatgactgag agccaggcta tggatattca 180
1363 aacatgtatt tttcagacac ttcttgaaaa taagtgaagc aaacctgtta attacaaggg 240
1364 aagcaatgac aatatttgtt gccaatgata a                          271
1814 <210> SEQ ID NO: 76
1815 <211> LENGTH: 247
1816 <212> TYPE: DNA
1817 <213> ORGANISM: Homo sapiens
1819 <400> SEQUENCE: 76
1820 tgtatattga gctcctactg tgtggcaagg cctatggtaa gcattttatt ttggtaactt 60
1821 gtttaatcct cattacaatt ctgtggtaaa tgctattatc tgttttttata ttgaagggat 120
E--> 1822 gaaatggagg ctcagaggga tatgtagtag ctaaatgtta gagctaggat tganacccaa 180
1823 attgacttct gagtatagat ttcccccccaa ctgtatgata cttcatattt ggagtcagct 240
1824 tgaagta                                                    247
2080 <210> SEQ ID NO: 91
2081 <211> LENGTH: 385
2082 <212> TYPE: DNA
2083 <213> ORGANISM: Homo sapiens
2085 <400> SEQUENCE: 91
2086 gaaatgggtc cggacaggtt aaaacaaaaa tccaatactg ccgtagtttc taggtggata 60
2087 taacattttt agaaatctta taatacaata ttaacttcat tggctgaacc caagcctttc 120
2088 agcctttata gatttgccat gatcctaata catataagca ttcattgtat tcattattaa 180
2089 ttacttcata gattcagtgt gtgacgaagg gagatgattt ttaacaaata ataaagtgaa 240
E--> 2090 atgatctagt tttgctatgt tgnttgagca acatcaaata gttttgctaa aatagataat 300
```

see item 9 on Error Summary Sheet

item 9

item 9

RAW SEQUENCE LISTING   DATE: 12/31/2001
PATENT APPLICATION: US/10/016,634   TIME: 14:11:01

Input Set : A:\Dex-255.app
Output Set: N:\CRF3\12312001\J016634.raw

item 9

```
E--> 2091 ttatagtgat ttttttttca ctatggnatt ttcttaaata tattaagggc tttcattttc 360
     2092 tgataccacc tagtttaatt ggggg                                       385
     2187 <210> SEQ ID NO: 94
     2188 <211> LENGTH: 1646
     2189 <212> TYPE: DNA
     2190 <213> ORGANISM: Homo sapiens
     2192 <400> SEQUENCE: 94                                          → item 9
E--> 2193 ttaggtccaa actcaggtna ncaaaactag tcccntcagn ccaattgctt gtacatttcn 60   → item 9
E--> 2194 acaggcaccn ctttggcaaa cccacatgga tnntnanaaa tggagaatga ggagacagcc 120  → item 9
E--> 2195 nnatattaat nagatgtatc aaactgtnac aatatgtgaa gagtattgtg tatatacaaa 180  → item 9
     2196 caggaaacaa ttgaaagcct tcaacatgtg tgggtggggg gagagataac tgaattaaca 240
     2197 ggccatgtag taaaacttaa aatcaaatcc agtagtcttg aaggtatagt aattgtttag 300
     2198 ttttgaaggt atagtaatta agtactgcgc actaaaaaaa actgaccaaa aggccgggtg 360
     2199 cggtggctca cgcctgtaat cccagcactt tgggaggccg aggcgggcgg atcacctgag 420
     2200 gtcaggagtt cgagaccagc ctggccaaca tggtgaaacc ccgtctctac taaaaacaca 480
     2201 aaaattagcc gggcatggtg gctcacgcct gtaatcccag cactctggga ggccaagaac 540
     2202 atgctgatgg tttatgccac aagttgacaa ctgtgtgtcc aactgtgaaa cctcagactc 600
     2203 aaggtctagc aaaagatgct tgggaaatcc ctcgagaatc tttgcgacta gaggttaaac 660
     2204 taggacaagg atgtttcggc gaagtgtgga tgggaacatg gaatggaacc acgaaagtag 720
     2205 caatcaaaac actaaaacca ggtacaatga tgccagaagc tttccttcaa gaagctcaga 780
     2206 taatgaaaaa attaagacat gataaacttg ttccactata tgctgttgtt tctgaagaac 840
     2207 caatttacat tgtcactgaa tttatgtcaa aaggtgctta ttccctttct attcgtgatt 900
     2208 gggatgagat aaggggtgac aatgtgaaac actacaaaat taggaaactt gacaatggtg 960
     2209 gatactatat cacaaccaga gcacaatttg atactctgca gaaattggtg aaacactaca 1020
     2210 cagaacatgc tgatggttta tgccacaagt tgacaactgt gtgtccaact gtgaaacctc 1080
     2211 agactcaagg tctagcaaaa gatgcttggg aaatccctcg agaatctttg cgactagagg 1140
     2212 ttaaactagg acaaggatgt ttcggcgaag tgtggatggg aacatggaat ggaaccacga 1200
     2213 aagtagcaat caaaacacta aaaccaggta caatgatgcc agaagctttc cttcaagaag 1260
     2214 ctcagataat gaaaaaatta agacatgata aacttgttcc actatatgct gttgtttctg 1320
     2215 aagaaccaat ttacattgtc actgaattta tgtcaaaaga gctccacaga cgctgcacag 1380
     2216 ctgtgaactc aactccagcc ctcagggaag gcagctcgca aagacctagg gcagcccgcc 1440
     2217 gatttcccaa aaagagcact gagaacccag caagcgtttt tctatatgat gctggcggca 1500
     2218 gcccagtgag ggaagtgcaa caagtatggt tctcctgctt gtcaactgtg gaaacagcga 1560
     2219 ccctgaaagt ggaggagcca cagctggaat cgtgttcgtc tgaatacacg cacctttcct 1620
     2220 atgagccttg taaagccagt cgttga                                     1646
     3517 <210> SEQ ID NO: 153
     3518 <211> LENGTH: 63
     3519 <212> TYPE: PRT
     3520 <213> ORGANISM: Homo sapiens
     3522 <400> SEQUENCE: 153
E--> 3523 Met Gln Cys Thr Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
     3524   1               5                  10                  15
E--> 3526 Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
     3527              20                  25                  30
E--> 3529 Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Lys Ile Lys Phe Gly
     3530           35                  40                  45
     3532 Met Lys Gln Glu Leu Ser Trp Thr Ile Tyr Asn Leu Leu Arg Tyr
     3533        50                  55                  60
```

item 9

VERIFICATION SUMMARY   DATE: 12/31/2001
PATENT APPLICATION: US/10/016,634   TIME: 14:11:02

Input Set : **A:\Dex-255.app**
Output Set: **N:\CRF3\12312001\J016634.raw**

```
L:13 M:270 C: Current Application Number differs, Replaced Application Number
L:14 M:271 C: Current Filing Date differs, Replaced Current Filing Date
L:100 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:2
L:101 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:2
L:102 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:2
L:103 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:2
L:163 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:6
L:164 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:6
L:165 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:6
L:382 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:16
L:412 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:17
L:436 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:18
L:552 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:22
L:701 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:25
L:703 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:25
L:726 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:26
L:727 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:26
L:769 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:28
L:829 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:30
L:830 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:30
L:831 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:30
L:832 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:30
L:1087 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:42
L:1163 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:46
L:1361 M:340 E: (46) "n" or "Xaa" used: Feature required, for SEQ ID#:56
L:1531 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:65
L:1532 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:65
L:1533 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:65
L:1551 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:65
L:1598 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:67
L:1599 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:67
L:1621 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:68
L:1622 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:68
L:1623 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:68
L:1698 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:71
L:1740 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:72
L:1741 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:72
L:1784 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:73
L:1822 M:340 E: (46) "n" or "Xaa" used: Feature required, for SEQ ID#:76
L:1884 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:79
L:1885 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:79
L:1886 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:79
L:1906 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:80
L:1907 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:80
L:1954 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:83
L:1955 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:83
L:1956 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:83
L:1957 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:83
```

VERIFICATION SUMMARY
PATENT APPLICATION: US/10/016,634

DATE: 12/31/2001
TIME: 14:11:02

Input Set : **A:\Dex-255.app**
Output Set: **N:\CRF3\12312001\J016634.raw**

```
L:1958 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:83
L:1959 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:83
L:1960 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:83
L:1961 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:83
L:2090 M:340 E: (46) "n" or "Xaa" used: Feature required, for SEQ ID#:91
M:340 Repeated in SeqNo=91
L:2178 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:93
L:2179 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:93
L:2193 M:340 E: (46) "n" or "Xaa" used: Feature required, for SEQ ID#:94
M:340 Repeated in SeqNo=94
L:2591 M:258 W: Mandatory Feature missing, <223> not found for SEQ ID#:113
L:2594 M:258 W: Mandatory Feature missing, <223> not found for SEQ ID#:113
L:2692 M:258 W: Mandatory Feature missing, <223> not found for SEQ ID#:117
L:2778 M:258 W: Mandatory Feature missing, <223> not found for SEQ ID#:121
L:2781 M:258 W: Mandatory Feature missing, <223> not found for SEQ ID#:121
L:2798 M:258 W: Mandatory Feature missing, <223> not found for SEQ ID#:122
L:2843 M:258 W: Mandatory Feature missing, <223> not found for SEQ ID#:124
L:2846 M:258 W: Mandatory Feature missing, <223> not found for SEQ ID#:124
L:2849 M:258 W: Mandatory Feature missing, <223> not found for SEQ ID#:124
L:3523 M:340 E: (46) "n" or "Xaa" used: Feature required, for SEQ ID#:153
M:340 Repeated in SeqNo=153
L:3565 M:258 W: Mandatory Feature missing, <223> not found for SEQ ID#:155
L:3600 M:258 W: Mandatory Feature missing, <223> not found for SEQ ID#:156
L:3603 M:258 W: Mandatory Feature missing, <223> not found for SEQ ID#:156
L:3655 M:258 W: Mandatory Feature missing, <223> not found for SEQ ID#:159
L:3694 M:258 W: Mandatory Feature missing, <223> not found for SEQ ID#:161
L:3718 M:258 W: Mandatory Feature missing, <223> not found for SEQ ID#:162
L:3721 M:258 W: Mandatory Feature missing, <223> not found for SEQ ID#:162
L:3756 M:258 W: Mandatory Feature missing, <223> not found for SEQ ID#:164
L:3759 M:258 W: Mandatory Feature missing, <223> not found for SEQ ID#:164
L:3762 M:258 W: Mandatory Feature missing, <223> not found for SEQ ID#:164
L:3765 M:258 W: Mandatory Feature missing, <223> not found for SEQ ID#:164
L:3768 M:258 W: Mandatory Feature missing, <223> not found for SEQ ID#:164
L:3771 M:258 W: Mandatory Feature missing, <223> not found for SEQ ID#:164
L:3774 M:258 W: Mandatory Feature missing, <223> not found for SEQ ID#:164
L:3777 M:258 W: Mandatory Feature missing, <223> not found for SEQ ID#:164
L:3891 M:258 W: Mandatory Feature missing, <223> not found for SEQ ID#:170
L:3960 M:258 W: Mandatory Feature missing, <223> not found for SEQ ID#:172
L:3963 M:258 W: Mandatory Feature missing, <223> not found for SEQ ID#:172
L:3966 M:258 W: Mandatory Feature missing, <223> not found for SEQ ID#:172
L:3980 M:258 W: Mandatory Feature missing, <223> not found for SEQ ID#:173
L:3983 M:258 W: Mandatory Feature missing, <223> not found for SEQ ID#:173
```